

# Multimodal Image Analysis in Alzheimer's Disease via Statistical Modelling of Non-local Intensity Correlations

Marco Lorenzi<sup>1</sup>, Ivor J. Simpson<sup>1</sup>, Alex F. Mendelson<sup>1</sup>, Sjoerd B.  
Vos<sup>1,2</sup>, M. Jorge Cardoso<sup>1</sup>, Marc Modat<sup>1</sup>, Jonathan M. Schott<sup>3</sup>,  
and Sebastien Ourselin<sup>1</sup>

<sup>1</sup>Translational Imaging Group, CMIC, UCL, London, UK

<sup>2</sup>MRI Unit, Epilepsy Society, Chalfont St Peter, UK

<sup>3</sup>Dementia Research Centre, UCL Institute of Neurology, Queen  
Square, London

Corresponding author:

Marco Lorenzi

3rd floor, Wolfson House, 4 Stephenson Way

University College London

London, NW1 2HE

United Kingdom

mail: m.lorenzi@ucl.ac.uk

Data used in preparation of this article were obtained from the Alzheimer's

Disease Neuroimaging Initiative (ADNI) database ([adni.loni.usc.edu](http://adni.loni.usc.edu)). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: [http://adni.loni.usc.edu/wp-content/uploads/how\\_to\\_apply/ADNI\\_Acknowledgement\\_List.pdf](http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf) (2014) (Date of access: 07/2015)

## Abstract

The joint analysis of brain atrophy measured with magnetic resonance imaging (MRI) and hypometabolism measured with positron emission tomography with fluorodeoxyglucose (FDG-PET) is of primary importance in developing models of pathological changes in Alzheimer’s disease (AD). Most of the current multimodal analyses in AD assume a *local* (spatially overlapping) relationship between MR and FDG-PET intensities. However, it is well known that atrophy and hypometabolism are prominent in different anatomical areas. The aim of this work is to describe the relationship between atrophy and hypometabolism by means of a data-driven statistical model of non-overlapping intensity correlations. For this purpose, FDG-PET and MRI signals are jointly analyzed through a computationally tractable formulation of partial least squares regression (PLSR). The PLSR model is estimated and validated on a large clinical cohort of 1049 individuals from the ADNI dataset. Results show that the proposed non-local analysis outperforms classical local approaches in terms of predictive accuracy while providing a plausible description of disease dynamics: early AD is characterised by non-overlapping temporal atrophy and temporo-parietal hypometabolism, while the later disease stages show overlapping brain atrophy and hypometabolism spread in temporal, parietal and cortical areas.

## Introduction

The multimodal analysis of anatomical and physiological images is of primary importance in developing comprehensive models of biological processes and pathologies, and increasing the statistical power of current imaging biomarkers. Already, both brain atrophy, measured in magnetic resonance images (MRIs), and hypometabolism, quantified by positron emission tomography with fluorodeoxyglucose radiotracers (FDG-PET), are among the primary diagnostic biomarkers of Alzheimer’s disease (AD). The information provided by these two imaging modalities is correlated, since hypometabolism and neuronal loss are interdependent biological phenomena. However, at the present moment, a joint model of the hypometabolism-atrophy relationship in AD has not been developed, and current hypotheses on their interaction are mostly based on the quantification of grey matter volume and FDG uptake at the regional level.

In recent years, voxel-wise approaches to multimodal analysis in AD have been proposed [1]. In particular, image synthesis techniques based on machine learning have been used to synthesise FDG-PET images from MRIs of AD patients for diagnostic purposes [2, 3]. The majority of these approaches are based on the local modelling of the relationship between MR and FDG-PET signals, either by considering the voxels independently, or through neighborhoods (patches) defined around voxels. However, it is well known that the link between morphology and function in the brain is not purely local [4]. For this reason, local methods may provide only a limited description of the link between structure and function in AD.

Several techniques have been proposed for modelling non-overlapping signal correlations in the field of functional MRI analysis. For instance, both independent component analysis (ICA) or partial least square (PLS) approaches have been successfully applied to the joint analysis of functional activation in

the brain and covariates drawn from genetic, clinical, or imaging data [5, 6, 7]. In the context of correlation modelling in multimodal imaging analysis, multivariate techniques such as PLS have the appealing characteristic that they do not rely on any hypothesis about the spatial overlap between voxels' signals. They are thus able to model relationships between non-adjacent voxels. Unlike purely local correlation model approaches (for example, those based on voxel-wise correspondencies or on patch-based search windows), these methods optimise the *latent components* describing the *global* correlation of the images treated as multidimensional arrays. This enables them to model the potentially significant interactions between voxels located in completely different areas of a single image, or between voxels in images of different modalities.

Several multivariate approaches have previously been applied to the multimodal analysis of imaging data in neurodegenerative diseases [8]. Notable approaches include parallel ICA, which has been used to analyze the relationships between brain amyloid deposition and either atrophy or hypometabolism [9, 10], and canonical correlation analysis, which has been used to study the correlation between structural connectivity and brain atrophy [11]. Though PLS itself has been previously applied in the joint analysis of brain metabolism and atrophy [12, 13], past analyses have been limited to relatively small clinical samples and have focused on solely the first latent component. The use of higher-order components may aid in the discovery of more complex correlation structures, though it brings with it greater challenges related to stability and replicability.

The aim of this work is to investigate the spatial relationship between brain atrophy and hypometabolism in a large clinical cohort of the ADNI dataset, by means of a data-driven PLS statistical model of non-overlapping intensity correlations. This is achieved by applying a computationally tractable formulation

of PLS regression (PLSR) to the joint analysis of non-local intensity correlations in FDG-PET and T1 weighted MR images. Unlike previous studies, in this work we extend PLSR to the analysis of the high-order latent components, and we introduce a thorough cross-validation scheme in order to identify the reproducible and biologically relevant latent components of joint correlation.

The performance of the PLSR model is compared to a common non-parametric approach for multimodal image analysis based on local intensity similarities. The experimental validation shows that the proposed PLSR approach outperforms the local reference analysis in terms of predictive accuracy while providing an interpretable, reproducible and biologically plausible description of the spatial relationship between atrophy and hypometabolism in AD.

## Local vs Non-local Correlation Models of Imaging Data

In this section, we introduce the computational models used to compare non-local and local assumptions of multimodal intensity correlation. The respective models are a computationally tractable application of PLSR to non-local intensity correlation in image data [14] and a non-parametric model based on local patch similarities [15, 16].

### Computationally Tractable PLSR in Imaging Data

In the following, let  $\mathbf{X} = (X_k)_{k=1}^{N_s}$  and  $\mathbf{Y} = (Y_k)_{k=1}^{N_s}$  be the matrices of predictor and predicted image modalities respectively, where  $\{X_k, Y_k\}$  is the multimodal image pair sampled at the same voxel grid, of subject  $k$ . We assume that  $\mathbf{X}$  and  $\mathbf{Y}$  correspond to T1-MR and FDG-PET respectively. The size of  $\mathbf{X}$  and  $\mathbf{Y}$  is  $N_s \times N$ , where  $N_s$  is the number of individuals,  $N$  is the number of image

voxels, and the images are represented by row vectors.

The partial least squares (PLS) approach is based on the decomposition of the observations through a projection onto  $m$ -dimensional latent spaces defined by the basis vectors  $\mathbf{T} = (\mathbf{t}_l)_{l=1}^m$  and  $\mathbf{U} = (\mathbf{u}_l)_{l=1}^m$  such that  $\mathbf{X} = \mathbf{TP}^T + \mathbf{E}$ , and  $\mathbf{Y} = \mathbf{UQ}^T + \mathbf{F}$ , where  $\mathbf{P}$  and  $\mathbf{Q}$  are the associated coefficients, and  $\mathbf{E}$  and  $\mathbf{F}$  are matrices of residuals. In particular, PLS aims to maximise the covariance of the projections in the latent space:  $\text{cov}(\mathbf{t}, \mathbf{u})^2 = \max_{\mathbf{w}, \mathbf{c}} [\text{cov}(\mathbf{X}\mathbf{w}, \mathbf{Y}\mathbf{c})]^2$ , where  $\mathbf{w}$  and  $\mathbf{c}$  are unitary basis vectors of the latent space. Several formulations of PLS have been proposed in different research contexts [17, 18, 19, 20, 21], and it can be shown that the solution of PLS can be obtained from the principal vectors of the singular value decomposition (SVD) of the covariance matrix  $\mathbf{X}^T\mathbf{Y}$  [18]. PLS can be iteratively computed as follows:

Let  $\mathbf{X}^{(0)} = \mathbf{X}$ ,  $\mathbf{Y}^{(0)} = \mathbf{Y}$ . Iterate over the index  $i = 1, \dots, m \leq N_s$ :

1. *SVD step.* Compute the principal eigen-vectors  $\mathbf{w}_i$  and  $\mathbf{c}_i$  of the SVD decomposition of the matrix  $\mathbf{R}_i = \mathbf{X}^{(i)T} \mathbf{Y}^{(i)}$ .
2. Compute the latent vectors  $\mathbf{t}_i = \mathbf{X}^{(i)} \mathbf{w}_i$ , and  $\mathbf{u}_i = \mathbf{Y}^{(i)} \mathbf{c}_i$ , and corresponding coefficients  $\mathbf{p}_i = \mathbf{X}^{(i)T} \mathbf{t}_i$ .
3. *Deflation step.* Decorrelate the data from the principal directions:  $\mathbf{X}^{(i+1)} = \mathbf{X}^{(i)} - \mathbf{t}_i \mathbf{p}_i^T$ , and  $\mathbf{Y}^{(i+1)} = \mathbf{Y}^{(i)} - \mathbf{u}_i \mathbf{c}_i^T$ .

We note that the matrix  $\mathbf{R}_i$  is usually very large (voxels $\times$ voxels), and its SVD decomposition is generally computationally infeasible. However, the SVD step can still be efficiently computed from the eigen-value problem associated with the matrix  $\mathbf{X}\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T$ , which is usually of much smaller dimension ( $N_s \times N_s$ ). This approach has been proposed previously in [14], which focused on the analysis of within-modality non-local intensity correlations in neuroimages. In

particular, this efficient optimization scheme was used to model either group-wise patterns of cortical thickness from MRI, or functional connectivity networks measured in fMRI. In this work, we apply this computational approach in the context of multimodal analysis of brain images.

PLSR builds upon the above formulation of PLS by assuming a linear relationship between the vectors  $\mathbf{t}$  and  $\mathbf{u}$ , i.e.  $\mathbf{U} = \mathbf{TB} + \mathbf{H}$ , where  $\mathbf{B}$  is a latent linear mapping and  $\mathbf{H}$  is the residual matrix. The PLS model can thus be rewritten as  $\mathbf{Y} = \mathbf{TC}^T + \mathbf{F}^*$ , where  $\mathbf{C}^T = \mathbf{BQ}^T$ , and  $\mathbf{F}^* = \mathbf{HQ}^T + \mathbf{F}$  is the residual error. It can be shown that the solution of the PLSR is  $\tilde{Y} = X\mathbf{D}_{PLS}$  where the regression coefficient  $\mathbf{D}_{PLS} = (\mathbf{P}^T)^+\mathbf{BQ}^T$ , and  $(\mathbf{P}^T)^+ = \mathbf{P}(\mathbf{P}^T\mathbf{P})^{-1}$  is the right Moore-Penrose pseudoinverse of  $\mathbf{P}^T$  [20].

## PLSR model of the non-overlapping spatial correlations in multimodal images

PLSR has a number of appealing features that can be exploited in the context of high-dimensional modelling of multimodal images. First, the basis of spatial eigen-components,  $\{\mathbf{w}_i, \mathbf{c}_i\}$ , provides a *parsimonious* and low-dimensional representation of the multimodal correlation pattern, and can be used for exploratory analysis and modelling purposes. Second, the loadings  $(X^T \mathbf{w}_i)_{i=1}^m$  are a low-dimensional representation of the individual anatomy that can be used to address quantitative analysis problems, such as group-wise comparisons or classification. Finally, PLSR defines a transfer function linking the two modalities and, given an individual image  $X$ , provides a prediction of the associated target modality  $Y$  according to the model of non-local signal correlation estimated in the data  $(\mathbf{X}, \mathbf{Y})$ .

### **Relationship between PLSR and principal component regression.**

Another intuitive approach to multimodal correlation modelling in image data is principal component regression (PCR). This consists of an initial PCA step applied separately to the predicted and predictor variables, and a subsequent regression step to model the correlations between the resulting low dimensional representations. Both PCR and PLSR yield a predictive model using low dimensional latent space, but they differ in how this is driven; while PLSR aims to find a subspace that directly maximises the covariance between the predicted and predictor variables, PCR selects a latent space that maximises the variability within each variable set separately. For this reason, it may include components that are not useful in characterising the relationship between variable sets which may degrade predictive performance. We therefore prefer PLSR.

The interested reader can find an experimental illustration of the differences between PLSR and PCR when applied to the problem analysed in this study in Appendix A.

### **Local Models of Intensity Correlations Based on Patch Similarities**

Patch-based methods are becoming a popular approach for the estimation of non-linear signal correspondences between different modalities. They have found several successful applications in medical image analysis, such as multi-modal image registration [15], or in FDG-PET synthesis [2]. This approach is based on the assumption that, given an individual image  $X$ , the intensities of the target modality  $\tilde{Y}$  can be inferred from the local intensity correspondences observed in a database of atlas pairs  $\{X_j^A, Y_j^A\}_{j=1}^{N_A}$  [22]. The between-modality voxel-to-voxel mapping  $X(s) \rightarrow \tilde{Y}(s)$  is usually not explicitly modelled in a parametric way, but is indirectly inferred from the local intensities of the images  $Y_j^A(s)$

corresponding to the atlases  $X_j^A$  most correlated with  $X$ . The local correlation model presented here is the same as the one proposed in state-of-art approaches in the context of FDG-PET image synthesis [16]. All subjects images are aligned to the target using non-linear registration, and the intensity at a given location is estimated using the most similar patch in the database as determined using the local intensity information (the  $L^2$  metric). The chosen patch size was of 5 voxels.

## **Analysis of Brain Hypometabolism and Atrophy in Alzheimer’s Disease**

### **Study Participants**

Data used in the preparation of this article were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database ([adni.loni.usc.edu](http://adni.loni.usc.edu), Date of Access: 04/2013). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging, positron emission tomography, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment and early Alzheimer’s disease.

Patients were selected when both MR and FDG-PET images were available at the baseline timepoint. The resulting study cohort consisted of 1049 subjects: 274 healthy controls, 168 patients affected by AD, and 607 patients with mild cognitive impairment (MCI). Of the latter, 154 subsequently converted to AD during the time of the study. Clinical and socio-demographical information are reported in Table 1.

	healthy	MCI stable	MCI conv	AD
N	274	453	154	168
age (years)	74.1 (5.98)	72.16 (7.55)	73.21 (7.37)	75.66 (7.66)
sex (% females)	48	42	41	42
education (years)	16.22 (2.77)	16.01 (2.73)	16.03 (2.67)	15.1 (3.08)
MMSE	28.98 (1.21)	28.11 (1.66)	27.02 (1.75)	23.05 (2.1)

Table 1: Baseline socio-demographical and clinical information of the cohort of this study. The entries for age, education and MMSE indicate group-wise mean and standard deviation in parenthesis. MMSE: mini mental state exam.

## Image Processing

FDG-PET images were obtained at the standardised resolution of 8 mm FWHM, and normalised using the mean intensity in the cerebellar grey matter. T1-weighted MR images at both 1.5 and 3.0 Tesla were included to increase the size of the available sample. A sample specific group-wise space was defined for our analysis using iterative non-rigid registration and averaging the grey matter segmented from the MR images. Registration was performed using the freely available nifty-reg package [23], and grey matter and FDG-PET images were resampled to the group-wise space. The resampled grey matter images were modulated by the corresponding Jacobian determinant of the template-to-subject transformation, and subsequently spatially filtered at the point spread function of the PET images and downsampled. Thanks to the modulation and to the downsampling operation, the resulting anatomical areas of apparent ventricular and CSF expansion are associated to smaller voxel-wise intensity values due to the scaling by the Jacobian determinant values.

## Statistical Analysis

### Model Estimation and Comparison

The goodness of fit of PLSR and local approaches was assessed by cross-validation. The training data was composed of 80 healthy controls, 80 MCI, and 80 AD pa-

tients, randomly chosen from the study cohort. The data were respectively used to 1) estimate the PLSR latent components and regression coefficients, and 2) as an atlas database for the local patch-based method (PM). The PLSR model was computed by estimating 30 latent components. The model could be ideally computed by estimating the 239 latent basis-components corresponding to  $N - 1$  training data samples. The number of estimated PLSR components was limited to 30 for practical reasons because, as shown in the next experimental section, the stability and reproducibility of the high-order components is generally very low, and they usually provide very little contribution to the model performance. The resulting multimodal correlation models were validated on the remaining subjects. The experiment was repeated 10 times with different training sets to ensure the generalisation of the results. Due to the non-parametric nature of PM, we also compared the PM trained with a leave-one-out scheme (1048 training samples per test) in order to use of the largest amount of training data. The predictive accuracy was measured by the absolute difference between predicted and observed FDG-PET in temporal, posterior and parietal cortices, and by comparing the average predicted FDG-PET regional values to the SUVR values independently reported in the ADNI dataset. Figure 1 illustrates the flowchart of the cross-validation scheme of a single repeat adopted in the proposed experimental setting.

### **Reproducibility and Biological Plausibility of PLSR model**

We investigated the biological plausibility of the non-local correlation pattern of the PLSR model. For this purpose, a linear discriminant analysis (LDA) was performed on the coefficients of the latent space associated with the testing subjects, in order to identify the mostly discriminative PLS components through leave-one-out. The discriminative accuracy of the PLSR model was quantified by computing the area under the receiver operating characteristic (ROC) curve

## Model estimation and comparison for a single repeat

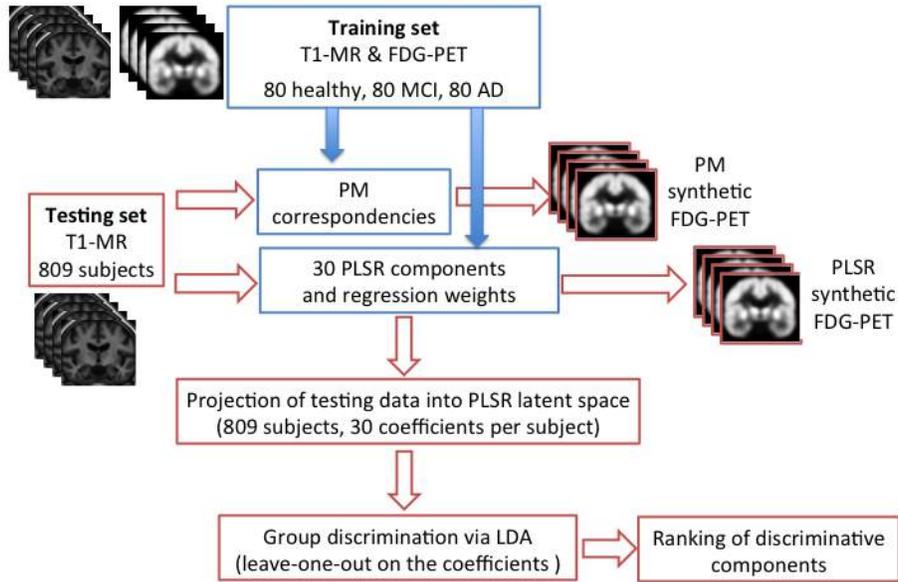


Figure 1: Flowchart of the cross-validation scheme of a single repeat in the proposed experimental setting.

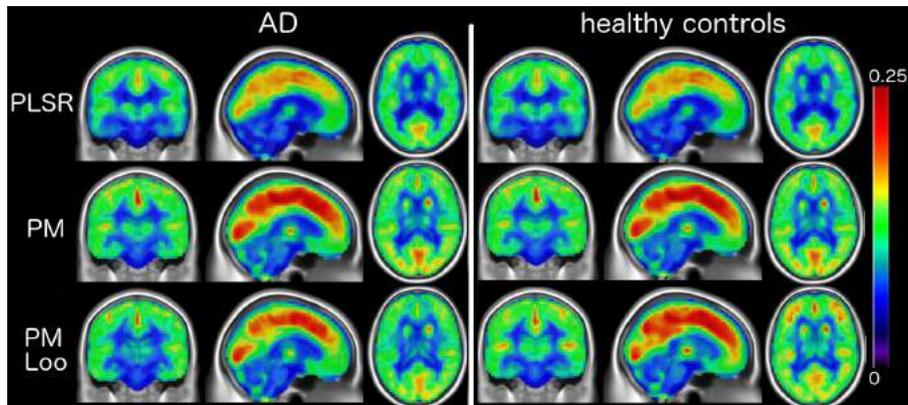


Figure 2: Mean absolute prediction error of PLSR and PM (240 training samples, and leave-one-out -Loo-). PLSR provides higher predictive accuracy than the local patch based (PM) approach. Results are similar when considering the MCI group (not shown).

associated to the LDA classification result.

The interpretation of the PLSR modes of correlation is usually challenging, since some of the obtained components (especially at the high order) tend to be noisy and not necessarily related to meaningful anatomical interpretation (an illustrative example of the set of the first 10 components estimated by PLSR in a single repeat is shown in Appendix B). For this purpose, in order to address this important issue of robustness, we measured the reproducibility of the most discriminative components across the 10 repeats. We were interested primarily in the rate of reproduction of the individual components, rather than in their relative ordering. To this end, the discriminative power of each component was quantified by the absolute value of the associated LDA weight, and the resulting 5 mostly discriminative components of each experiment were matched to those of the other repeats. Components were matched when the absolute value of the correlation between them exceeded 0.5. When multiple matches were possible, the strongest was chosen. The resulting labelling establishes the reproducibility of the discriminative components across repeats, and does not necessarily reflect the order of the eigen-components estimated in each PLS run.

## Results

### Model comparison

Figure 2 shows the average pattern of absolute differences of the predictions in AD and healthy controls with both methods. We notice that PLSR generally provides a better fit, while the local patch based method leads to larger estimation errors in parietal and temporal areas. We note that increasing the training sample size slightly improves the PM, especially in the temporal regions. The average regional absolute error between predicted images and real ones was sys-

tematically higher for PM as compared to PLSR, and significantly different for 8 out of 10 repeats ( $p < 0.01$ , paired t-test). The PLSR prediction also provided significantly better agreement with the ADNI measurements than PM (Table 2A). This is reflected by the significantly higher effect size associated with the average measures for the PLSR approach, indicating a better separation between clinical groups (Table 2B).

### Biological Plausibility of the PLSR model

Figure 3, left, shows the reproducibility results for the PLSR components across repeats. The components are ordered according to the order of output in the PLSR results obtained in the first cross-validation repeat. The reproducibility results are instead quantified by the green and red bars, which indicates the number of times that a component was among the most discriminative across repeats. We note that only few of the mostly discriminative PLSR components are highly reproducible across repeats. In particular, the first eigen-component estimated in each PLSR repeat is the most discriminative when comparing AD vs healthy controls, while the third one is most discriminative when comparing both MCI stable and converters, and healthy controls and MCI converters. Both components were 100% reproducible across repeats.

A. Correlation wrt ADNI SUVR			B. Effect size		
	PLSR	PM		PLSR	PM
Whole cohort*	0.31 (0.25, 0.38)	0.21 (0.14, 0.27)	AD vs HC *	1.07	0.79
AD and HC*	0.33 (0.22, 0.43)	0.23 (0.12, 0.34)	MCIc vs MCIs	0.53	0.37
MCIc and MCIs*	0.30 (0.22, 0.38)	0.20 (0.12, 0.28)	MCIc vs HC*	0.67	0.46

Table 2: A. Correlation (mean, 95% confidence interval) between predicted average regional FDG-PET and the corresponding SUVR values reported in ADNI. B. Effect size between the measures obtained with PLSR and with PM. HC: healthy controls, MCIc: MCI converted to AD, MCIs: MCI stable. (\* for significant differences,  $p < 0.05$ , paired t-test)

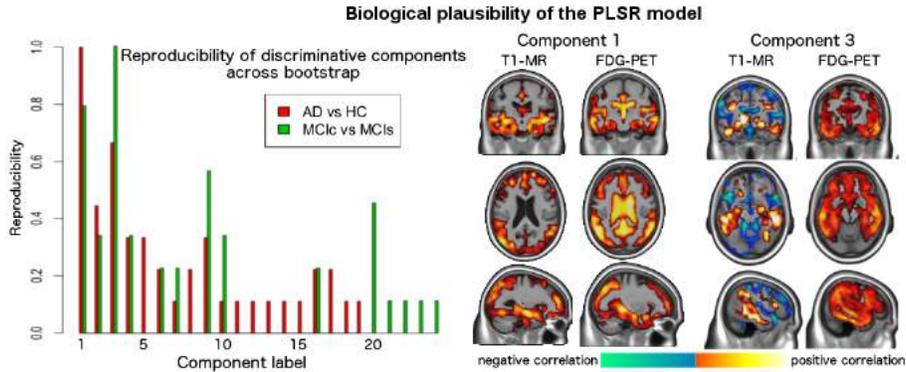


Figure 3: Left: reproducibility of the discriminative components. When comparing AD and controls, component 1 was the only one 100% reproducible and discriminative across repeats. The same consideration applies to component 3 when comparing stable and converting MCI. Right: component 1 describes the relationship between atrophy and FDG-PET uptake spread in temporal, parietal and posterior regions. We also note the partial volume effect in the ventricles for the FDG component. Component 3 shows the *non-overlapping* spatial inverse relationship between increased expansion of the CSF (ventricles and brain sulci), and joint increased temporal atrophy and cortical hypometabolism.

We note that since the figure is relative to the first cross validation repeat, the components estimated during the other repeats that did not find any match in the reproducibility analysis are omitted in the figure. However, apart from the reported components 1 and 3, no other reproducible component associated to the other repeats was observed in the analysis.

These components are shown on the right hand side of Figure 3, and the associated correlation network is shown in Figure 4. A 3D rendering of the correlation networks is shown as the following url: [https://www.dropbox.com/s/orsf3nt6hq2kp38/supplementary\\_animations1.mov?dl=0](https://www.dropbox.com/s/orsf3nt6hq2kp38/supplementary_animations1.mov?dl=0) (12/10/2015). The networks were obtained by thresholding components 1 and 3, and by subsequently applying a morphological opening operation in order to identify a consistent set of clusters of maximal PLSR weights.

On one hand, we note that component 1 describes the relationship between

## AD-related networks of joint variation between atrophy and hypometabolism

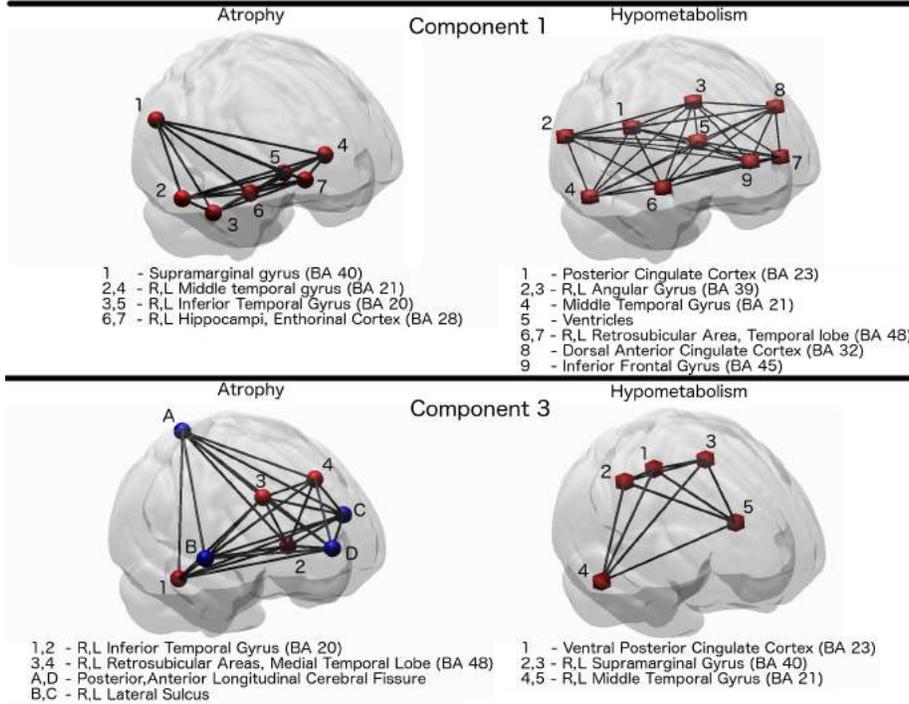


Figure 4: PLSR networks of joint relationship between atrophy and hypometabolism derived from Component 1 and Component 3. Red hubs indicate regions of joint within- and between-modality correlation. Blue hubs indicate anticorrelated regions (CSF expansion measured in T1-MR). BA = Broadmann anatomical areas.

atrophy and FDG-PET uptake spread in temporal, parietal and posterior regions. In particular, it shows the partial volume effect due to ventricular expansion in AD, that is already observable in the raw data, and that leads to the very large variation of the FDG-PET signal in the ventricles in subjects with pronounced global brain atrophy. On the other hand, component 3 shows the non-overlapping inverse correlation pattern between increased expansion of the CSF, and joint increased temporal atrophy and cortical hypometabolism.

Finally, the average area under the ROC curve for the classification tasks across the different folds was 0.87 (0.83, 0.91 - 95% c.i.) for the comparison

of AD vs healthy controls, and 0.75 (0.73, 0.76 - 95% c.i.) for the comparison between stable and converting MCI. This result primarily confirms the ability of the proposed PLSR to model biologically relevant features, and is in line with the classification performance based on T1-MR information previously reported in the literature on the ADNI dataset [24, 25, 26].

## Conclusions

We have investigated the problem of multimodal analysis of biomedical images in AD, by comparing two different modelling hypothesis based on state-of-art techniques, PLSR and patch-based local correlation, to promote non-local correlation analysis approaches with respect to localized ones in describing multimodal correlation patterns in AD. Our study introduces and validates the use of PLSR in the context of multimodal modelling in AD by showing that PLSR bases and coefficients can be estimated in very large datasets of volumetric images through a computationally tractable approach to the eigen-decomposition.

### **Non-local vs local multimodal modelling in AD.**

Our results show that the proposed non-local approach outperforms classical PM-based multimodal local correlation models in terms of modelling accuracy and predictive power. The ensemble of the reported results proves the ability of the proposed PLSR in capturing biologically relevant features, and in generalising to unseen structural imaging data of T1-MR scans.

Even though the presented study does not provide a theoretical proof of the superiority of non-local methods, our results show that T1-MR and FDG-PET present reproducible and consistent patterns of correlations between non-overlapping anatomical areas. This study thus shows that realistic multimodal models of neurodegeneration necessarily need to account for the non-local rela-

tions intimately related to the neurobiological aspects of the disease.

### **Plausibility of the PLSR model.**

PLSR provides a parsimonious description of the global biological variability, represented by the low-dimensional latent subspace parameterisation. For this reason, the interpretation and statistical analysis of PLSR is more straightforward than that of the usually complex models provided by non-parametric local approaches [2].

Our analysis revealed that in the sequential stages of the pathology (cognitively normal  $\rightarrow$  MCIc  $\rightarrow$  AD) we can consistently identify two reproducible components of correlation between atrophy and hypometabolism. Our results are therefore supportive of the existence of different patterns of atrophy and hypometabolism which differentially characterise the different stages of the disease, and thus are informative of the dynamics of the pathology.

The correlation networks highlighted in this study are supported by known biological dynamics between atrophy and hypometabolism in dementia [4]: although hypometabolism and atrophy are typically locally correlated, i.e. areas with neuronal loss (atrophy) show by definition reduced or absent metabolism, hypometabolism may be seen in areas not obviously or typically affected by atrophy, as exemplified by focal dementia syndromes, such as posterior cortical atrophy.

This work shows that T1 weighted MRI and FDG-PET in AD are highly correlated and share important patterns of common non-overlapping spatial relationship. The proposed method could be used in the future to identify and decorrelate the common inter-modality variation from biomedical images for the identification of more specific image based biomarkers.

## Acknowledgments

Sebastien Ourselin receives funding from the EPSRC (EP/H046410/1, EP/J020990/1, EP/K005278), the MRC (MR/J01107X/1), the EU-FP7 project VPH-DARE@IT (FP7-ICT-2011-9-601055), the NIHR Biomedical Research Unit (Dementia) at UCL and the National Institute for Health Research University College London Hospitals Biomedical Research Centre (NIHR BRC UCLH/UCL High Impact Initiative-BW.mn.BRC10269). Jonathan M. Schott acknowledges the support of the NIHR UCL/H Biomedical Research Centre, the NIHR Queen Square Biomedical Research Unit in Dementia, and grant funding from the MRC, Alzheimer’s Research UK, and Leonard Wolfson Experimental Neurology Centre

Data collection and sharing for this project was funded by the Alzheimer’s Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer’s Association; Alzheimer’s Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the

Foundation for the National Institutes of Health ([www.fnih.org](http://www.fnih.org)). The grantee organisation is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer’s Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for NeuroImaging at the University of Southern California.

## Additional Information

The author(s) declare no competing financial interests.

## References

- [1] Frisoni, G.B., et al. In vivo mapping of amyloid toxicity in Alzheimer disease. *Neurology*, 72:17, 1504-11, 2009.
- [2] Rongjian, L., et al. Deep learning based imaging data completion for improved brain disease diagnosis. *Proc. of Medical Image Computing and Computer Assisted Intervention (MICCAI)*, LNCS, Springer, 8675, 305-312, 2014.
- [3] Burgos, N. et al. Subject-specific Models for the Analysis of Pathological FDG PET Data. *Proc. of Medical Image Computing and Computer Assisted Intervention (MICCAI)*, LNCS, Springer, 9351, 651-658, 2015.
- [4] Jack, C.R., et al. Hypothetical model of dynamic biomarkers of the Alzheimer’s pathological cascade. *Lancet Neurol*, 9:1, 119-128, 2010.
- [5] Krishnan, A. et al. Partial Least Squares (PLS) methods for neuroimaging: a tutorial and review. *Neuroimage*, 56:2,455–475, 2011.
- [6] Vergara, V.M., et al. A three-way parallel ICA approach to analyze links among genetics, brain structure and brain function. *NeuroImage*, 98, 386-394, 2014.
- [7] Burzynska, A.Z., et al. A scaffold for efficiency in the human brain, *J Neurosci*, 33:43, 17150-17159, 2013.

- [8] Sui, J., Adalib, T., Yu, Q., Calhoun, V.D. A review of multivariate methods for multimodal fusion of brain imaging data. *J Neurosci. Methods*, 204:1, 68-81, 2012.
- [9] Tosun, D., et al. *Spatial patterns of brain amyloid- $\beta$  burden and atrophy rate associations in mild cognitive impairment*, *Brain*, 134 (4): 1077-1088, 2011.
- [10] Laforce, R., et al. Parallel ICA of FDG-PET and PiB-PET in three conditions with underlying Alzheimer's pathology. *NeuroImage: Clinical*, 4, 508-516.
- [11] Avants, B, et al. Dementia induces correlated reductions in white matter integrity and cortical thickness: a multivariate neuroimaging study with sparse canonical correlation analysis, *Neuroimage*, 50(3), 1004-1016, 2010.
- [12] Chen, K., et al. Linking functional and structural brain images with multivariate network analyses: a novel application of the partial least square method, *NeuroImage*, 47:2, 602-610, 2009.
- [13] Chen, K., et al. Correlations between FDG PET glucose uptake-MRI gray matter volume scores and apolipoprotein E  $\epsilon$ 4 gene dose in cognitively normal adults: a cross-validation study using voxel-based multi-modal partial least squares, *NeuroImage*, 60(4), 2316-2322, 2012.
- [14] Worsley, K.J., Chen, J., Lerch, J., and Evans, A.C. Comparing Functional Connectivity via Thresholding Correlations and Singular Value Decomposition. *Phil Trans: Biological Sciences*, 360:1457, 913-920, 2005,
- [15] Iglesias, J.E., et al. Is Synthesizing MRI Contrast Useful for Inter-modality Analysis? *Proc. of Medical Image Computing and Computer Assisted Intervention (MICCAI)*, LNCS, Springer, 8149, 631-638, 2013.
- [16] Burgos, N., et al. Attenuation Correction Synthesis for Hybrid PET-MR Scanners: Application to Brain Studies. *IEEE TMI*, 33:12, 2232-42, 2014.
- [17] Wold, H., et al. Estimation of principal components and related models by iterative least squares. *Multivariate analysis*, 1: 391-420, 1966.
- [18] Agnar, H. PLS regression methods. *J Chemometrics*, 2, 3, John Wiley & Sons, Ltd., 1988.

- [19] Geladi, Paul and Kowalski, Bruce R. Partial least-squares regression: a tutorial, *Analytica chimica acta*, 185: 1–17, 1986.
- [20] Rosipal, R. and Krämer, N. Overview and recent advances in partial least squares. *Subspace, Latent Structure and Feature Selection Techniques*, LNCS, Springer, 34–51, 2006.
- [21] Vinzi, V., et al. *Handbook of partial least squares*, Springer, 1-22, 2010.
- [22] Hertzmann, A., et al. *Image analogies*, Proceedings of the 28th annual conference on Computer graphics and interactive techniques (SIGGRAPH), Los Angeles, 327–340. New York: ACM, 2001.
- [23] Modat, M., et al. Fast free-form deformation using graphics processing units. *Comput Meth Prog Bio*, 98:3, 278-284. <http://sourceforge.net/projects/niftyreg> (18/09/2009) (Date of access: 03/2014)
- [24] Chincarini, A., et al. Local MRI analysis approach in the diagnosis of early and prodromal Alzheimer’s disease. *NeuroImage*, 58(2):469-80, 2011.
- [25] Young, J., et al. Accurate multimodal probabilistic prediction of conversion to Alzheimer’s disease in patients with mild cognitive impairment. *NeuroImage: Clinical*, 2, 735:745, 2013.
- [26] Davatzikos, C., Bhatt, P., Shaw, L.M., Batmanghelich, K.M., Trojanowski, J.Q. Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. *Neurobiol Aging*, 32:12, 2322.e19-2322.e27, 2011.

## 1 Authors contribution

"M. L." designed and implemented the PLSR code, performed the statistical analysis and wrote the main manuscript text. "I. J. S." and "M. J. C." provided feedback on the statistical content, and helped in the designing of the local correlation technique of section 2.3. "A. F. M." and "M. M." worked on the preprocessing of the ADNI data used in the experimental section 4.1. "S.B.V." provided support in the

conception and visualization of the networks of Figure 4. "J. S." and "S. O." supervised the project, and provided feedback on the analysis and on the interpretation of the results. All authors reviewed the manuscript.